Matrices Satisfying the Van Der Waerden Conjecture

SHMUEL FRIEDLAND
The Weizmann Institute of Science
Rehovot, Israel

Recommended by Hans Schneider

ABSTRACT

It is shown here that any $n \times n$ doubly stochastic matrix whose numerical range lies in the sector from $-\pi/2n$ to $\pi/2n$ satisfies the van der Waerden conjecture.

1. INTRODUCTION

Let A be an $n \times n$ doubly stochastic matrix. An unresolved conjecture of van der Waerden [5] states that

$$p(A) \geqslant \frac{n!}{n^n}. \tag{1.1}$$

Let $\mathcal{P}_r$ be the collection of n-square complex matrices whose numerical range lies in the sector from $-\pi/2r$ to $\pi/2r$. The inequality (1.1) was proved by Marcus and Newman [3] for a symmetric nonnegative definite A, i.e., $A \in \mathcal{P}_\infty$. Afterwards, Sasser and Slater [4] showed (1.1) for a normal doubly stochastic matrix belonging to the set $\mathcal{P}_n$. In this paper we prove the van der Waerden conjecture for any doubly stochastic matrix belonging to $\mathcal{P}_n$.

2. PRELIMINARIES

Let $Q = (q_{\mu\nu})$ be an $m \times n$ complex valued matrix. By Q' and Q^* we denote the transpose and the conjugate transpose respectively. The rth induced matrix $P_r(Q)$ is defined as follows [2, p. 20]. Denote by $G_{k,n}$ the totality of nondecreasing sequences of k integers chosen from $\{1, \ldots, n\}$. Let $\alpha \in G_{k,n}$. The integer $\mu(\alpha)$ is defined to be the product of the factorials of the multiplicities of the distinct integers appearing in the sequence α. Now $P_r(Q)$ is the $\binom{m+r-1}{r}\times\binom{n+r-1}{r}$ matrix whose entries are $p(Q[\alpha|\beta])/\sqrt{\mu(\alpha)\mu(\beta)}$ arranged lexicographically in $\alpha=(\alpha_1,\dots,\alpha_r)$ and $\beta=(\beta_1,\dots,\beta_r)$. Recall that $p(Q[\alpha|\beta])$ is the permanent of the matrix $(q_{\alpha_\mu\beta_\nu})_1^r$. If S is an $n\times p$ matrix then $P_r(QS)=P_r(Q)P_r(S)$. By $A\geqslant 0(>0)$ we mean that A is a hermitian nonnegative (positive) definite matrix. Denote by $C\geqslant B$ if C-$B\geqslant 0$ and $C>B$ if C-$B>0$. Let $u=(u_1,\dots,u_n)$ be a nontrivial complex valued column vector. A matrix $J(u)=(h_{\mu\nu}(u))_1^n$ is defined as follows

$$J(u)u=u,\qquad J(u)v=0\quad \text{if } v^*u=0. \tag{2.1}$$

Clearly $J(u)\geqslant 0$. Note that $J(u)$ is similar to a diagonal matrix D_n of the form

$$D_n=\operatorname{diag}\{1,0,\dots,0\}. \tag{2.2}$$

The converse of this statement is true:

LEMMA 1. *Let A be a hermitian matrix similar to D_n. Then $A=J(u)$ for some u and the vector u is unique up to a multiplicative constant.*

Proof. The similarity of A to D_n implies that 1 is a simple root of the characteristic polynomial and other eigenvalues vanish. Thus u is the unique eigenvector corresponding to the eigenvalue 1. As A is hermitian from the condition $v^*u=0$ it follows that v is a linear combination of the remaining n-1 eigenvectors. To these eigenvectors corresponds the eigenvalue zero. Therefore $Av=0$ and $A=J(u)$ by definition.

Denote by $\mathfrak{M}_n(u)$ the set of all $n\times n$ complex valued matrices which satisfy

$$Au=A^*u=u. \tag{2.3}$$

LEMMA 2. *Let A belong to the set $\mathfrak{M}_n(u)$. If $A+A^*\geqslant 0$ then*

$$\frac{A+A^*}{2}\geqslant J(u). \tag{2.4}$$

Proof. Clearly $B=(A+A^*)/2\in\mathfrak{M}_n(u)$. Furthermore $(B$-$J(u))u=0$ and $v^*(B$-$J(u))v\geqslant 0$ if $v^*u=0$. Thus B-$J(u)\geqslant 0$.

COROLLARY 1. *Let A belong to the set $\mathfrak{M}_n(u)$. If $P_r(A)+P_r(A^*)\geqslant 0$ then*

$$\frac{P_r(A)+P_r(A^*)}{2}\geqslant P_r(J(u)). \tag{2.5}$$

Proof. First note that $P_r(A)\in \mathfrak{M}_{\binom{n+r-1}{r}}(P_r(u))$. The matrix $P_r(J(u))$ is similar to $D_{\binom{n+r-1}{r}}$. Furthermore (2.1) implies that

$$P_r(J(u))P_r(u)=P_r(u).$$

Therefore by Lemma 1 $P_r(J(u))=J(P_r(u))$. Now (2.5) follows by Lemma 2.

The condition (2.5) implies many inequalities involving the permanents of the submatrices of A and $J(u)$. As an application we shall consider some special cases. Let $e=(1,\dots,1)$. Then all entries of $J(e)$ are equal to $1/n$.

COROLLARY 2. *Let A belong to $\mathfrak{M}_n(e)$. If $P_r(A)+P_r(A^*)\geqslant 0$ then*

$$\operatorname{Re}\{p(A[\alpha|\alpha])\}\geqslant \frac{r!}{n^r}, \tag{2.6}$$

for any $\alpha\in G_{r,n}$.

Proof. By Corollary 1, $B=P_r(A)+P_r(A^*)-2P_r(J(e))\geqslant 0$. In particular the diagonal entries of B are nonnegative which means the inequality (2.6).

3. CHARACTERIZATION OF A CERTAIN CLASS OF MATRICES

Let $\mathcal{P}_r$ and $\mathcal{P}_r^0$ be the collection of n-square complex matrices whose numerical range lies in the closed and open sectors from $-\pi/2r$ to $\pi/2r$, i.e.,

$$\mathcal{P}_r=\{A|\operatorname{Re}\{x^*Ax\}\geqslant \operatorname{ctg}\pi/2r|\operatorname{Im}\{x^*Ax\}|\}, \tag{3.1a}$$

$$\mathcal{P}_r^0=\{A|\operatorname{Re}\{x^*Ax\}> \operatorname{ctg}\pi/2r|\operatorname{Im}\{x^*Ax\}|\}, \tag{3.1b}$$

where $r\geqslant 1$. It is easy to see that an equivalent definition of the sets $\mathcal{P}_r$ and $\mathcal{P}_r^0$ is

$$\mathcal{P}_r=\{A|(A+A^*)\pm i\operatorname{ctg}\pi/2r(A-A^*)\geqslant 0\}, \tag{3.1c}$$

$$\mathcal{P}_r^0=\{A|(A+A^*)\pm i\operatorname{ctg}\pi/2r(A-A^*)> 0\}. \tag{3.1d}$$

The following theorem characterizes those matrices for which $P_r(A)+P_r(A^*)>0$.

THEOREM 1. *The following statements are equivalent*

(1) $P_r(A)+P_r(A^*)>0$.

(2) *There exists an rth root of the unity* $\omega^r=1$ *such that the matrix* ωA *belongs to the set* $\mathcal{P}_r^0$.

Proof. (1)→(2). Let C be a $k\times k$ complex valued matrix. By $\mathcal{D}(C)$ we denote the following domain in the complex plane.

$$\mathcal{D}(C)=\{z/z=x^*Cx, x^*x=1\}. \tag{3.2}$$

Note that $\mathcal{D}(C)$ is a compact connected domain. It is easy to see that $C+C^*>0)$ iff $\mathcal{D}(C)$ is contained in the open right half plane $H=\{z|\ \operatorname{Re}\{z\}>0\}$. From the equality $P_r(x^*)P_r(C)P_r(x)=(x^*Cx)^r$ it follows that $\mathcal{D}^r(C)\subset\mathcal{D}(P_r(C))$ where $\mathcal{D}^r(C)=\{z/z=\zeta^r, \zeta\in\mathcal{D}(C)\}$. Thus the assumption $P_r(A)+P_r(A^*)>0$ implies that

$$\mathcal{D}^r(A)\subset H. \tag{3.3}$$

This shows that the domain $\mathcal{D}(A)$ does not contain the origin. Combining the relation (3.3) with the connectivity of $\mathcal{D}(A)$ we see that the domain $\mathcal{D}(A)$ is contained in the domain

$$G=\left\{z/|\arg(z\omega)|<\frac{\pi}{2r}\right\}$$

for some rth root of unity ω. Thus for any nontrivial vector x the following inequality holds:

$$\operatorname{Re}\{x^*(\omega A)x\}>\operatorname{ctg}\frac{\pi}{2r}|\{\operatorname{Im}x^*(\omega A)x\}|.$$

This is exactly the statement (2).

(2)→(1). Let $B=\omega A+\bar{\omega}A^*$ and $C=i(\omega A-\bar{\omega}A^*)$.

We first show that there exists a nonsingular matrix X and a diagonal matrix $D=\operatorname{diag}\{d_1,\ldots,d_n\}$ such that

$$\omega A=X^*DX, \tag{3.4}$$

and

$$|\arg d_\nu|<\pi/2r, \qquad \nu=1,\ldots,n. \tag{3.5}$$

Indeed, since the sum of two positive definite matrices is positive definite, we conclude from (3.1d) that $B>0$. In that case it is well known that the two hermitian matrices B and C can be simultaneously diagonalized [1, p. 313]. That is, there exists a nonsingular matrix X such that

$$B = X^* D_1 X, \qquad C = X^* D_2 X,$$

where D_1 and D_2 are two real diagonal matrices

$$D_1 = \operatorname{diag}\{d_1^{(1)}, \dots, d_n^{(1)}\}, \qquad D_2 = \operatorname{diag}\{d_1^{(2)}, \dots, d_n^{(2)}\},$$

and $d_\nu^{(1)} > 0$, $\nu = 1, \dots, n$. Moreover as $B \pm \operatorname{ctg}(\pi/2r)C > 0$ we obtain that

$$\operatorname{tg}\frac{\pi}{2r} > \frac{|d_\nu^{(2)}|}{d_\nu^{(1)}}, \qquad \nu = 1, \dots, n. \tag{3.6}$$

Thus $\omega A = X^* D X$ where $D = (D_1 - iD_2)/2$ and the inequalities (3.6) imply (3.5). Now

$$P_r(\omega A) + P_r(\bar{\omega} A^*) = P_r(X^*)[P_r(D) + P_r(D^*)]P_r(X).$$

But $P_r(D)$ is also a diagonal matrix whose diagonal entries are of the form

$$\prod_{j=1}^{r} d_{\nu_j}, \qquad 1 \leqslant \nu_1 \leqslant \nu_2 \leqslant \cdots \leqslant \nu_r \leqslant n.$$

By (3.5) we have

$$|\arg \prod_{j=1}^{r} d_{\nu_j}| < \frac{\pi}{2}.$$

Therefore $P_r(D) + P_r(D^*) > 0$ and $P_r(\omega A) + P_r(\bar{\omega} A^*) > 0$. We note that if $\omega^r = 1$ then

$$P_r(A) = P_r(\omega A), \qquad P_r(A^*) = P_r(\bar{\omega} A^*)$$

This shows that $P_r(A) + P_r(A^*) > 0$. The proof of the theorem is completed.

Note that if A belongs to $\mathfrak{M}_n(u)$ then $u^* u = u^* A u$ and the point $z = 1$ lies in $\mathfrak{D}(A)$. Therefore if $P_r(A) + P_r(A^*) > 0$ from the proof of Theorem 1 it follows that A itself belongs to $\mathfrak{P}_r^0$. From Theorem 1 we deduce the following corollary.

COROLLARY 3. *Assume that A belongs to the set $\mathfrak{P}_r$. Then*

$$P_r(A) + P_r(A^*) \geqslant 0. \tag{3.7}$$

Proof. Let $A_\epsilon = A + \epsilon I$ where $\epsilon > 0$ and I is the identity matrix. Since the sum of nonnegative definite matrix with positive definite one is positive definite the assumption $A\epsilon\mathcal{P}_r$ (3.1c) implies that $A_\epsilon\epsilon\mathcal{P}_r^0$ (3.1d). By Theorem 1

$$P_r(A_\epsilon)+P_r(A_\epsilon{}^*)>0 \text{ and letting } \epsilon\to 0 \text{ we obtain (3.7).}$$

REMARK. It seems that the condition (3.7) implies that ωA belongs to $\mathcal{P}_r$ where $\omega^r=1$. However it does not follow only from the fact that $\mathcal{D}^r(A)$ contained in the closed right half plane as it was in the proof of Theorem 1.

4. INEQUALITIES IN PERMANENTS

Combining Corollary 3 with Corollary 2 we obtain the following Theorem.

THEOREM 2. *Let A be an $n\times n$ doubly stochastic matrix. If A belongs to the set $\mathcal{P}_r$ then*

$$p(A[\alpha|\alpha])\geqslant\frac{r!}{n^r}, \tag{4.1}$$

for any $\alpha\epsilon G_{r,n}$.

Suppose now that A is a symmetric doubly stochastic matrix which is nonnegative definite. Clearly A belongs to any set $\mathcal{P}_r$ and we obtain by Theorem 2 the inequalities (4.1) for arbitrary positive integer r. This extends the result demonstrated in [3]. We consider the equality sign in (4.1).

THEOREM 3. *Let $A=(a_{ij})_1^n$ be an $n\times n$ doubly stochastic nonnegative definite matrix. Then for any natural number r and $\alpha\epsilon G_{r,n}$ $p(A[\alpha|\alpha])\geqslant r!/n^r$. If for some $\alpha=(\alpha_1,\dots,\alpha_r)$ the equality sign holds, then*

$$a_{k\alpha_i}=a_{\alpha_i k}=\frac{1}{n}, \qquad i=1,\dots,r, \quad k=1,\dots,n. \tag{4.2}$$

Proof. By Corollaries 1 and 3 $P_r(A)-P_r(J(e))\geqslant 0$. Thus if $p(A[\alpha|\alpha])-(r!/n^r)=0$ then all entries of $P_r(A)-P_r(J(e))$ on the α row vanish. So

$$p(A[\alpha|\beta])=\frac{r!}{n^r}, \tag{4.3}$$

for any $\beta\epsilon G_{r,n}$. Take $\beta=(k,\dots,k)$. Now

$$\prod_{i=1}^{r}a_{\alpha_i k}=\frac{1}{n^r}. \tag{4.4}$$

Let w_i be the multiplicity of i in the vector $(\alpha_1,\ldots,\alpha_r)$, $i=1,\ldots,n$. Note that

$$\prod_{i=1}^{r} a_{\alpha_i k} = \prod_{i=1}^{n} (a_{ik})^{w_i}.$$

By the arithmetic-geometric mean inequality

$$\frac{1}{n} = \left(\prod_{i=1}^{n} (a_{ik})^{w_i} \right)^{1/r} \leqslant \sum_{i=1}^{n} \frac{w_i}{r} a_{ik}, \qquad k=1,\ldots,n.$$

Finally

$$1 \leqslant \sum_{k=1}^{n} \sum_{i=1}^{n} \frac{w_i}{r} a_{ik} = \sum_{i=1}^{n} \frac{w_i}{r} = 1,$$

which means that $a_{\alpha_i k} = 1/n$, $i=1,\ldots,r$, $k=1,\ldots,n$.

We suspect that for any doubly stochastic matrix A belonging to $\mathcal{P}_r$ the equality $p(A[\alpha|\alpha]) = r!/n^r$ implies (4.2). For the permanent of A we have the following theorem.

THEOREM 4. *Let A be an* $n \times n$ *doubly stochastic matrix belonging to the set* $\mathcal{P}_n$. *Then*

$$p(A) > p(J(e)), \tag{4.5}$$

if $A \neq J(e)$.

Proof. By Corollary 3, $P_n(A) + P_n(A') \geqslant 0$. Thus Corollary 1 implies $P_n(A) + P_n(A') - 2P_n(J(e)) \geqslant 0$. If $p(A) - p(J(e)) = 0$ then

$$p(A[\alpha|\beta]) + p(A[\beta|\alpha]) = \frac{2n!}{n^n},$$

for any $\beta \epsilon G_{n,n}$ and $\alpha = (1,\ldots,n)$. Take $\beta = (k,\ldots,k)$. Now

$$\prod_{i=1}^{n} a_{ik} + \prod_{i=1}^{n} a_{ki} = \frac{2}{n^n}. \tag{4.6}$$

But

$$\prod_{i=1}^{n} a_{ik} \leqslant \frac{1}{n^n} \left(\sum_{i=1}^{n} a_{ik} \right)^n = \frac{1}{n^n},$$

$$\prod_{i=1}^{n} a_{ki} \leqslant \frac{1}{n^n} \left(\sum_{i=1}^{n} a_{ki} \right)^n = \frac{1}{n^n}.$$

Combining this with (4.6) we obtain that $A = J(e)$.

Assume now that A is a normal matrix. Then the numerical range of A is the convex hull of its eigenvalues. Thus A belongs to $\mathcal{P}_n$ if and only if the eigenvalues of A lie in the sector from $-\pi/2n$ to $\pi/2n$. For a normal doubly stochastic matrix belonging to the $\mathcal{P}_n$ the van der Waerden permanent conjecture was proved by Sasser and Slater [4]. Other inequalities for the sum of the permanents of all r-square submatrices of A appearing in [4] can be easily obtained by Corollary 1.

The author wishes to thank the referee for his valuable remarks.

Note added in proof. Recalling that the numerical range of A is a convex set and using the same arguments as in the proof of Theorem 1 we can deduce that the condition (3.7) implies $\omega A \epsilon P_r$.

REFERENCES

1 F.R. Gantmacher, *Theory of Matrices*, Vol. I, Chelsea, New York (1964).
2 M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber and Schmidt (1964).
3 M. Marcus and M. Newman, Inequalities for the permanent function, *Ann. Math.* **75** (1962), 47–62.
4 D.W. Sasser and M.L. Slater, On the inequality $\Sigma x_i y_i \geqslant 1/n \; \Sigma x_i \cdot \Sigma y_i$ and the van der Waerden permanent conjecture, *J. Combinatorial Theor.* **3** (1967), 25–33.
5 B.L. van der Waerden, Aufgabe 45, *Iber Deutsch. Math-Verein*, **35** (1926), p. 117.

Received February 16, 1973; revised July 31, 1973